Exposure to Mimivirus collagen promotes arthritis

Shah, Nikunj ; Hülsmeier, Andreas J ; Hochhold, Nina ; Neidhart, Michel ; Gay, Steffen ; Hennet, Thierry

Abstract: Collagens, the most abundant proteins in animals, also occur in some recently described nucleocytoplasmic large DNA viruses such as Mimiviridae, which replicate in amoebae. To clarify the impact of viral collagens on the immune response of animals exposed to Mimiviridae, we have investigated the localization of collagens in Acanthamoeba polyphaga mimivirus particles and the response of mice to immunization with mimivirus particles. Using protein biotinylation, we have first shown that viral collagen encoded by the ORF L71 is present at the surface of mimivirus particles. Exposure to mimivirus collagens elicited the production of anti-collagen antibodies in DBA/1 mice immunized intradermally with mimivirus protein extracts. This antibody response also targeted mouse collagen type II and was accompanied by T-cell reactivity to collagen and joint inflammation as observed in collagen-induced arthritis following immunization of mice with bovine collagen type II. The broad distribution of nucleocytoplasmic large DNA viruses in the environment suggests that humans are constantly exposed to such large virus particles. A survey of blood sera from human healthy subjects and from rheumatoid arthritis patients indeed demonstrated that 30% of healthy subject and 36% of rheumatoid arthritis sera recognized the major mimivirus capsid protein L425. Moreover, whereas 6% of healthy subject sera recognized the mimivirus collagen protein L71, 22% of rheumatoid arthritis sera were positive for mimivirus L71. Accordingly, our study shows that environmental exposure to mimivirus represents a risk factor in triggering autoimmunity to collagens.

DOI: https://doi.org/10.1128/JVI.03141-13

Exposure to Mimivirus Collagen Promotes Arthritis

Nikunj Shah[a], Andreas J. Hülsmeier[a], Nina Hochhold[a], Michel Neidhart[b], Steffen Gay[b], Thierry Hennet[a#]

[a]Institute of Physiology and Zurich Center of Integrative Human Physiology, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland.

[b]Center of Experimental Rheumatology, University Hospital Zurich and Zurich Center of Integrative Human Physiology, Gloriastrasse 23, CH-8091 Zurich, Switzerland.

Corresponding author: Thierry Hennet, Institute of Physiology, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. Tel: +41-44-635 50 80.

Email: thennet@access.uzh.ch

Running title: Mimivirus collagen and arthritis

Word count abstract: 235

Word count text: 3956

Abstract

Collagens, the most abundant proteins in animals, also occur in some recently described nucleocytoplasmic large DNA viruses such as *Mimiviridae*, which replicate in amoebae. To clarify the impact of viral collagens on the immune response of animals exposed to *Mimiviridae*, we have investigated the localization of collagens in *Acanthamoeba polyphaga* mimivirus particles and the response of mice to immunization with mimivirus particles. Using protein biotinylation, we have first shown that viral collagen encoded by the ORF L71 is present at the surface of mimivirus particles. Exposure to mimivirus collagens elicited the production of anti-collagen antibodies in DBA/1 mice immunized intra-dermally with mimivirus protein extracts. This antibody response also targeted mouse collagen type II and was accompanied by T-cell reactivity to collagen and joint inflammation as observed in collagen-induced arthritis following immunization of mice with bovine collagen type II. The broad distribution of nucleocytoplasmic large DNA viruses in the environment suggests that humans are constantly exposed to such large virus particles. A survey of blood sera from human healthy subjects and from rheumatoid arthritis patients indeed demonstrated that 30% of healthy subject and 36% of rheumatoid arthritis sera recognized the major mimivirus capsid protein L425. Moreover, whereas 6% of healthy subject sera recognized the mimivirus collagen protein L71, 22% of rheumatoid arthritis sera were positive for mimivirus L71. Accordingly, our study shows that environmental exposure to mimivirus represents a risk factor in triggering autoimmunity to collagens.

Introduction

Nucleocytoplasmic large DNA viruses (NCLDVs) represent a growing group of giant viruses found in various types of aquatic environments (1). NCLDVs include *Poxviridae*, *Asfarviridae*, *Iridoviridae*, *Ascoviridae*, *Phycodnaviridae*, *Mimiviridae,* and *Marseilleviridae* (2). The *Paramecium bursaria* chlorella virus PBCV-1 was the first large DNA virus characterized at the molecular level and shown to harbor a complex genome of 330 kbp (3). But the largest NCLDVs described to date belong to *Mimiviridae,* which occur in fresh and saline environments and replicate within amoebae (4). *Acanthamoeba polyphaga* mimivirus was the first *Mimiviridae* isolated from a cooling water tower and characterized in 2003 (5). Other members of *Mimiviridae* include megavirus isolated from a marine environment (6), mamavirus (7), and moumouvirus (8). *Mimiviridae* feature large capsids exceeding 400 nm diameter and harbor large genomes of more than 1 Mbp. The genomes of NCLDVs encode structural proteins and enzymes usually not found in viruses, such as aminoacyl-tRNA synthetases, DNA repair enzymes, potassium ion channel, protein kinases and glycosyltransferases (5, 9, 10).

Interestingly, *Mimiviridae* also express multiple collagen genes during their infectious life cycle in amoebae. For example, mimivirus expresses seven collagen genes, namely L71, R196, R239, R240, R241, L668, L669, already by 6 h post infection (11). Even the virophage Sputnik includes two collagen genes among its predicted 21 open reading frames (12). The functional relevance of these collagens is however presently unknown. First analysis of mimivirus proteins indicated that collagen is hydroxylated like animal collagen (13). Cryo-electron microscopy and atomic force microscopy studies failed to reveal any collagen-like

63 structures in mimivirus (14, 15), although the dense fibers surrounding mimivirus capsids
64 have been suggested to represent cross-linked glycosylated collagen (14).
65 The ubiquitous distribution of NCLDVs in aquatic environments (16, 17) suggests that
66 humans are constantly exposed to such viruses. Mimivirus cannot replicate in animal cells
67 but can be internalized by phagocytosis by mouse and human macrophages (18). The uptake
68 of mimivirus particles by human macrophages potentially leads to virus antigen
69 presentation and thereby to the generation of antibodies against virus proteins. Considering
70 the structural similarity between animal and *Mimiviridae* collagens, we made the hypothesis
71 that antibodies generated against *Mimiviridae* collagens may cross-react with animal
72 collagens and thereby contribute to an autoimmune response to collagenous structures in
73 animals previously exposed to *Mimiviridae*. The present study provides evidence supporting
74 this hypothesis, showing that arthritis can be triggered in mice immunized with mimivirus
75 particles and by unraveling increased occurrence of antibodies against mimivirus collagen
76 in rheumatoid arthritis patients.

Materials and Methods

Ethics Statement - All mouse experiments were performed in compliance with the Swiss Animal Protection Ordinance and approved by the local veterinary authority (Kantonales Veterinäramt Zürich, Switzerland). The human sera tested in this study were a part of previously existing collection and the experimental protocol approved by the Kantonale Ethik-Kommission Zürich (KEK).

Giant virus infection and protein extraction - *Acanthamoeba polyphaga* and mimivirus were provided by Didier Raoult (CNRS UMR6020, Université de la Méditerranée, Marseille). Marseillevirus (2) was isolated from a water sample collected from the Lake of Zurich. Amoebae were routinely cultured as monolayer in PYG medium at 28°C as previously described (5). Mimivirus and marseillevirus were added to multiplicity MOI 10 to amoebae and newly formed virus were collected from the culture supernatant 2 days post infection. Virus particles were suspended in 0.5 M Tris-HCl, pH 8.5, 0.2% CHAPS, 2 mM TCEP, 6 M guanidine hydrochloride and incubated at 65°C for 10 min. After cooling to room temperature, iodoacetamide was added to a final concentration of 3 mM and further incubated at room temperature for 40 min. After adding DTT to a final concentration of 15 mM, protein extracts were centrifuged at room temperature at 17,000 x g and proteins in the supernatant were precipitated with 12% trichloroacetic acid.

Surface biotinylation of mimivirus proteins - Purified mimivirus particles were suspended in PBS and sulfo-NHS-biotin (Thermo Scientific, Waltham, MA, USA) was added to a final concentration of 1 mg/ml. Virions were rotated for 30 min at room temperature

and the reaction was quenched by adding equal volume of 100 mM glycine in PBS. Virions were pelleted and washed twice with 100 mM glycine in PBS and proteins were extracted with guanidine hydrochloride as described above. Extracts were diluted 10-fold in PBS, 0.1 % CHAPS, containing proteinase inhibitors (Calbiochem Proteinase Inhibitor Cocktail III, Merck Millipore, USA) and subjected to avidin cartridge purification (ABSciex, Framingham, Massachusetts, USA). The cartridge was successively washed with 500 µl PBS, 0.1% CHAPS; 1 ml of 650 mM NaCl in 20 mM phosphate buffer pH 7.2, 0.1% CHAPS; 1 ml of PBS, 0.1% CHAPS and 1 ml of 0.1% CHAPS in H_2O, and biotinylated proteins were eluted with 800 µl 0.4% trifluoroacetic acid, 0.1% CHAPS. Proteins were precipitated with trichloroacetic acid and subjected to SDS-PAGE. Individual protein bands were excised and subjected to in-gel tryptic digestion, as previously described (19, 20) followed by LC-MS protein identification. LC-MS data were analyzed using Mascot (Matrix Science, London, UK; version 2.3.02). Mascot was set up to search a Swissprot concatenated target-decoy database (2011.01.11, 1049100 entries) assuming the digestion enzyme trypsin. Mascot was searched with a fragment ion mass tolerance of 0.60 Da and a parent ion tolerance of 10.0 PPM. Iodoacetamide derivative of cysteine was specified in Mascot as a fixed modification. Oxidation of methionine and biotinylation of lysine were specified in Mascot as variable modifications. Scaffold (version Scaffold_3.4.9, Proteome Software Inc., Portland, OR, USA) was used to statistically validate MS/MS based peptide and protein identifications.

Collagen induced arthritis model - DBA/1 mice were purchased from Charles River (Germany), bred and maintained in the animal facility of Institute of Physiology, University of Zurich. All experiments were performed in compliance with the Swiss Animal Protection

Ordinance and approved by the local veterinary authority (Kantonales Veterinäramt Zürich, Switzerland). Collagen induced arthritis was established as described before (21). Briefly, 6-8 week old mice were immunized intra-dermally in the tail either with PBS, bovine collagen type II (Chondrex, USA), mimivirus L71 collagen-like protein, mimivirus protein extract, or marseillevirus protein extract emulsified in Complete Freund's Adjuvant (CFA, Chondrex, USA). Each mouse received either 50 µl of PBS, 100-120 µg of bovine collagen type II, 150 µg of recombinant L71 collagen-like protein, 120-150 µg of mimivirus proteins, or 120-150 µg of marseillevirus proteins emulsified 1:1 in CFA in a total volume of 50 µl. Mice received 30 days later a booster injection of the same amount of antigen emulsified 1:1 in Incomplete Freund's Adjuvant (IFA, Chondrex, USA). Development of arthritis was monitored daily for 75 days post immunization. Severity was scored on a level of 0 (no inflammation) to 4 (most severe inflammation) per limb per mouse, thus allowing a maximum score of 16 per mouse (21).

Anti-collagen type II antibodies - Mouse blood sera were collected by heart puncture. Anti-mouse CII antibody titers were measured in blood sera by ELISA (Chondrex, USA) as per manufacturer's instructions.

Histology – Limbs were skinned and fixed overnight in 10% neutral buffered formalin. Tissues were further decalcified using Immunocal solution (Quartett, Germany) for 4-5 days, dehydrated and paraffin embedded. Sections of 5 µm were mounted on glass slides and stained with H&E.

Recall assay - Axillary, lateral axillary, superficial inguinal and popliteal lymph nodes from mice were collected 8-10 days post booster immunization. Aliquots of 100,000 cells in 100 µl complete RPMI-1640 medium were stimulated with antigens and incubated for 48 h in a CO_2 incubator at 37°C with 5% CO_2. Cells were stimulated in 100 µl either with medium alone as negative control or concanavalin-A (Sigma, Switzerland) at 3 µg/ml as positive control. T-cell proliferation grade denatured mouse collagen type II at 1 mg/ml (Chondrex, USA), T-cell proliferation grade denatured bovine collagen type II at 1 mg/ml (Chondrex, USA) and heat denatured mimivirus collagen L71 at 1.5 mg/ml were used as antigen. After 48 h, 1 µCi of [^{3}H]thymidine (Perkin-Elmer, USA) per well was added, incubated for 16-18 h and cells were harvested on 96-well glass filter (Perkin-Elmer, USA). Radioactivity was counted using a 96-well scintillation beta-counter (Wallac, Perkin-Elmer).

Anti-mimivirus ELISA - Mimivirus proteins were coated in microtiter plates at 0.1 µg per well in 100 µl PBS overnight at 4°C. Plates were washed thrice with PBS-0.05% Tween and blocked with PBS, 0.05% Tween, 1% bovine serum albumin at 37°C for 2 h. Plates were washed, 100 µl of diluted human and rabbit sera added and further incubated at room temperature for 1 h. After three wash steps, 100 µl of 1:5000-diluted biotinylated anti-human or anti-rabbit IgG antibody (BD Biosciences, Switzerland) were added for 2 h. Plates were washed and 100 µl of 1:1000-diluted streptavidin-HRP conjugate (BD Biosciences) added for 1 h in dark. Plates were washed, incubated for 2 min with 50 µl 3, 3' , 5, 5' tetramethylbenzidine (TMB) substrate (BD Biosciences) before stopping the reaction with 25 µl of 2 N H_2SO_4. Color development was measured at 440 nm.

Immunoprecipitation of mimivirus proteins – Aliquots of 25 µl of human sera were incubated with 30 µl of protein-G sepharose 4 Fast Flow (GE Healthcare, Switzerland) beads along with 80 µl PBS on a rotating shaker for 1 h at 4°C. After centrifugation at 500 x g for 5 min at 4°C, supernatants were discarded and beads incubated with 20 µg of mimivirus protein extract in 80 µl of PBS and further incubated on a rotating shaker for 30 min at 4°C. Beads were washed three times in PBS and antigen-antibody complexes were eluted from the beads by adding 40 µl of 0.1 M glycine, pH 2.7. After neutralization by addition of 20 µl of 1 M Tris-HCl pH 9, eluates were separated by SDS-PAGE. Slices of polyacrylamide gel excluding IgG chains were excised and subjected to in-gel tryptic digest as previously described (20) and peptides identified by tandem mass spectrometry as above.

Cloning, bacterial expression and purification – The mimivirus ORF L71 and L425 were custom synthesized (Genescipt, USA) and subcloned into the expression vector pET16b (Merck Millipore, Switzerland) linearized with *Xho*I and *Hind*III (for L71) or *Xho*I and *BamH*I (for L425). His_6-tagged recombinant proteins were expressed after transformation into *E. coli* BL21 (DE3) cells (Novagen, Switzerland) under induction of 0.2 mM IPTG at 32°C for 1.5 h. Recombinant proteins were purified over Ni-Sepharose 6 Fast Flow (GE Healthcare, Switzerland) gravity flow columns.

Western blotting – Aliquots of 15 µg of recombinant mimivirus L71 and L425 proteins were subjected to SDS-PAGE and transferred to PVDF membrane (Bio-Rad). A His_6-tagged 15 kDa fragment of the human GLT25D2 protein (22) was used as negative control. A 36 kDa His_6-tagged fragment of human collagen type III encompassing 114 [G-X-Y] repeats and lacking

192 N- and C-propeptides was provided by Christoph Rutschmann, Institute of Physiology,
193 University of Zurich. Blots were blocked in 1% polyvinylpyrrolidone (Sigma, Switzerland) +
194 5% dry milk solution overnight at 4°C, then washed three times for 5 min with TBS, 0.1%
195 Tween and incubated with human sera diluted 1:4000 for 2 h at room temperature. After
196 washing four times for 5 min, blots were incubated with anti-human IgG-HRP (Promega,
197 Switzerland) at 1:7500 dilution at room temperature for 1 h. Blots were developed with
198 SuperSignal chemiluminescent substrate (Thermo Scientific).
199
200 **Statistical analysis** - One way ANOVA with Dunnett's multiple comparison (GraphPad
201 Prism) was performed to compare experimental groups.

Results

Surface localization of mimivirus collagen

To assess the possible localization of collagens at the surface of mimivirus, we have applied a biotinylation approach on mimivirus particles. Biotinylated mimivirus proteins were captured on streptavidin beads, eluted and identified by tandem mass spectrometry, which revealed 60 surface proteins including the collagen protein L71 (Table 1). The L71 protein has 945 amino acids with 4 collagen domains encompassing 561 amino acids (Fig. 1). The first collagenous domain of L71 includes a stretch with 73% sequence identity to a major human collagen type II T-cell epitope identified in rheumatoid arthritis (23, 24). Other identified surface proteins comprised the capsid protein L425, the putative GMC-type oxidoreductase R135, the thioredoxin domain-containing protein R362 among several uncharacterized proteins (Table 1).

Mimivirus proteins promotes arthritis in mice

Considering the surface expression of collagen L71, we have addressed the potential of mimivirus and recombinant L71 protein to induce joint inflammation in DBA/1 mice using the standard protocol for collagen-induced arthritis (21), which closely resembles rheumatoid arthritis in humans. Bovine collagen type II was used as positive control to induce arthritis. A protein extract of the giant virus marseillevirus, which lacks collagen-like proteins, was used as negative control. Intradermal immunization of bovine collagen type II and mimivirus protein extracts lead to joint inflammation as assessed by visual inspection and histological examination of limb tissues. Mice immunized with mimivirus proteins reached clinical scores of 6 whereas those immunized with bovine collagen type II reached 12 by 75 days (Fig. 2A). By contrast, immunization with recombinant L71 protein alone and

immunization with marseillevirus proteins failed to elicit joint inflammation (Fig. 2A). Altered cartilage integrity and synovial hyperplasia were evident in joints of mice immunized with bovine collagen type II and to a lesser extent in mice immunized with mimivirus proteins (Fig. 2B).

The breakage of immune tolerance induced by bovine collagen and mimivirus proteins was confirmed by detecting elevated serum titers of anti-mouse collagen type II IgG in mice immunized with bovine collagen type II and mimivirus proteins, whereas mice immunized with PBS and marseillevirus did not show elevated anti-mouse collagen type II IgG titers (Fig. 3). Mice immunized with recombinant L71 protein also showed significantly elevated anti-mouse collagen type II IgG titers, yet by an order of magnitude lower than the titers observed in mice immunized with mimivirus proteins (Fig. 3). The lower antibody response achieved with L71 immunization may account for the lack of joint inflammation seen in these mice. The cross-reactivity of T-cells was investigated in recall assays (25). Cells isolated from draining lymph nodes of the immunized mice were found to proliferate in response to *in vitro* presentation of denatured fragments of mouse collagen type II (Fig. 4A), bovine collagen type II (Fig. 4B), and mimivirus collagen protein L71 (Fig. 4C), thereby confirming the presence of auto-reactive T-cells after immunization with mimivirus proteins. The proliferative response to collagen was strongest for cells isolated from mice immunized with mimivirus proteins. This finding was surprising since mice immunized with bovine collagen type II showed the highest score for limb inflammation and for anti-collagen IgG titers. The strong proliferative response of T-cells from mimivirus proteins-immunized mice may reflect the higher antigenicity of mimivirus collagen considering its peptide sequence divergence from mammalian collagen sequences.

Immunity to mimivirus in humans

To determine whether humans are commonly exposed to mimivirus, we first examined the presence of antibodies against mimivirus in 100 healthy subjects by ELISA using whole mimivirus proteins as antigens. Reactivity to mimivirus proteins was variable; 58 human sera showed significant IgG titers in the 5% range of titers observed in the sera of rabbits previously immunized with mimivirus proteins (Fig. 5). To identify the major mimivirus proteins recognized by human sera, we coupled the IgG fraction of sera from healthy subjects and rheumatoid arthritis patients to protein-G sepharose beads, which were further incubated with preparations of mimivirus proteins. Mimivirus proteins retained on the IgG-protein G beads were identified by mass spectrometric peptide sequencing after trypsin digestion. The major capsid protein L425 was found in all samples followed by the putative GMC type oxidoreductase R135 and core protein L410, which were found in seven from ten samples (Table 2). Interestingly, the most frequent mimivirus proteins recognized by human sera were surface proteins according to our surface biotinylation study (Table 1). Mimivirus collagens did not appear among the proteins recognized by sera. This absence may be related to the abundance of lysine in mimivirus collagens, thereby yielding very short tryptic peptides that remained below the detection range of mass spectrometric peptide sequencing. The recognition of multiple mimivirus proteins by human sera confirmed the exposure of humans to mimivirus.

To further validate the occurrence of antibodies against specific mimivirus proteins in human sera were analyzed by Western blot. The major capsid L425 and collagen L71 proteins were expressed as His_6-tagged recombinant proteins in *E. coli* and purified on Ni^{2+}-sepharose columns. Pools of 100 healthy subject sera and 100 rheumatoid arthritis sera

were probed against the recombinant L425 and L71 mimivirus proteins. We examined the reactivity of sera towards surface collagen L71 since this protein was not detected among the mimivirus proteins captured by immobilized serum IgG in our previous experiment. For the 100 healthy subject and 100 rheumatoid arthritis sera tested, respectively 30 and 36 sera recognized the capsid L425 protein (Fig. 6A). This results confirmed that exposure to mimivirus is common in the human population. The detection of IgG against the mimivirus capsid protein L425 in 30% of tested sera suggests repeated antigenic challenge probably caused by repeated contact with mimivirus. Reactivity of human sera towards mimivirus collagen was more discriminatory. Whereas only 6 healthy subject sera recognized the mimivirus collagen L71, 22 rheumatoid arthritis sera were positive for the mimivirus collagen L71 (Fig. 6B). To exclude unspecific cross-reactivity of human sera towards polypeptides containing $[G\text{-}X\text{-}Y]_n$ collagen domains, we have tested the recognition of L71-positive sera for a fragment of human collagen type III encompassing 114 [G-X-Y] repeats and lacking N- and C-propeptides. None of the 28 human sera positive for mimivirus L71 did recognize the 36 kDa $[G\text{-}X\text{-}Y]_{114}$ construct (Fig. 7), thereby demonstrating the specificity of the antibody response to mimivirus L71 collagen. Accordingly, this work confirmed that the reactivity to mimivirus collagen was 3.5 times more frequent in the pool of rheumatoid arthritis sera in comparison to the limited reactivity of sera from healthy subjects.

Discussion

The present study demonstrated that mice generated auto-reactive anti-collagen antibodies after immunization with mimivirus proteins including viral collagens. A possible relationship between exposure to mimivirus collagen and the development of auto-immunity was corroborated by the occurrence of IgG against mimivirus collagen among rheumatoid arthritis patients. These findings suggested that repeated exposure to mimivirus leads to antibody formation to virus collagen and to a breakage of immune tolerance for endogenous collagens. Giant viruses like mimivirus are ubiquitous in the environment (16, 17), thereby supporting the frequent contact of humans to such viruses. Mimivirus is most likely ingested by water uptake and captured by dendritic cells and macrophages lining the gastrointestinal mucosa. Alternatively, virus particles may enter the airways as aerosol and be taken up by alveolar macrophages. In fact, mimivirus can be phagocytized by human and mouse macrophages although the virus cannot replicate in these cells (18). Along this line, mimivirus infection has been related to pneumonia in isolated cases, although without evidence for virus particles in disease cases (26, 27). This putative pathogenicity however does not preclude a more general effect of mimivirus on priming an auto-immune response. The sequence similarity between a stretch of mimivirus L71 and human collagen type II (Fig. 1) supports a possible cross-reactivity of antibodies due to antigenic mimicry. This was indeed confirmed in our study by the detection of anti-mouse collagen type II IgG in mice immunized with recombinant L71 protein. A similar case of antigenic mimicry occurs in *Campylobacter jejuni* infection, which causes gastroenteritis, but can lead to Guillain-Barré syndrome when antibodies against *Campylobacter* lipooligosaccharides cross-react with endogenous GM1 gangliosides on nerves cells (28). Likewise, the detection of antibodies

312 towards mimivirus L425 capsid protein in some *Francisella tularensis* infected patients
313 suggested cross-reactivity of mimivirus antigens with other microorganisms (29). But
314 surprisingly, no reactivity to mimivirus antigens was found in sera from healthy subjects in
315 the study of Pelletier *et al.* (29).
316 The main factors involved in the pathogenicity of rheumatoid arthritis could either be
317 genetic or environmental factors. Rheumatoid arthritis is an autoimmune disease with a
318 significant environmental component as supported by twin studies (30, 31). Repeated
319 contact to collagen antigens found in the environment may promote the development of
320 cross-reactive anti-collagen antibodies and to inflammation in collagen-rich tissues.
321 Antibodies against collagens can either recognize the triple helical conformation, or peptides
322 sequences in the triple helical domain or in telopeptides. The specificity of antibodies against
323 collagen depends on the activation of a humoral response alone, or on a combination of cell-
324 mediated and humoral responses (32). The reactivity of human sera to mimivirus collagen
325 L71 shown by Western blotting indicates that epitopes based on amino acid sequence and
326 not 3D conformations are being recognized. This notion was supported by finding no
327 reactivity of the L71-positive human sera for the human collagen type III $[G\text{-}X\text{-}Y]_{114}$
328 polypeptide used as negative control. The lack of recognition for the $[G\text{-}X\text{-}Y]_{114}$ polypeptide
329 also indicated that the reactivity towards L71 was specific to mimivirus exposure and not
330 the result of cross-reactivity to collagen-domain containing proteins such as those found in
331 some Gram-positive bacteria. We did find that mimivirus L71 protein was immunogenic and
332 lead to the production of anti-mouse collagen type II IgG, but immunization of DBA/1 mice
333 with recombinant L71 protein failed to induce arthritis. This recombinant protein likely did
334 not mimic the native conformation of collagen-like proteins, which is essential for

arthritogenicity. In fact, denaturation of collagen prior to immunization abrogates the arthritic response (33).

Altogether, in view of the structural similarity between mimivirus collagen and human collagen sequences, we propose that giant viruses expressing collagen represent a potential environmental risk factor contributing to the development of rheumatoid arthritis. A systematic survey of mimivirus distribution in the environment will contribute to a better appreciation of the environmental risk associated to such giant viruses in relation to the geographical incidence of rheumatoid arthritis.

Acknowledgements

We are grateful to Lubor Borsig, Daniel Rhyner, Carole Oertli, Jürg Cabalzar, Christoph Rutschmann and Charlotte Burger for technical assistance. We are grateful to the Functional Genomics Center Zurich for its support with mass spectrometric analysis. This work was supported by a Swiss National Science Foundation Grant 310030-129633 to TH.

References

1. **Wilson WH, Van Etten JL, Allen MJ.** 2009. The Phycodnaviridae: the story of how tiny giants rule the world. Curr Top Microbiol Immunol **328:**1-42.
2. **Colson P, Pagnier I, Yoosuf N, Fournous G, La Scola B, Raoult D.** 2013. "Marseilleviridae", a new family of giant viruses infecting amoebae. Arch Virol **158:**915-920.
3. **Van Etten JL, Burbank DE, Kuczmarski D, Meints RH.** 1983. Virus infection of culturable chlorella-like algae and dlevelopment of a plaque assay. Science **219:**994-996.
4. **Moliner C, Fournier PE, Raoult D.** 2010. Genome analysis of microorganisms living in amoebae reveals a melting pot of evolution. FEMS Microbiol Rev **34:**281-294.
5. **Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, La Scola B, Suzan M, Claverie JM.** 2004. The 1.2-megabase genome sequence of Mimivirus. Science **306:**1344-1350.
6. **Arslan D, Legendre M, Seltzer V, Abergel C, Claverie JM.** 2011. Distant Mimivirus relative with a larger genome highlights the fundamental features of *Megaviridae*. Proc Natl Acad Sci U S A **108:**17486-17491.
7. **Colson P, Yutin N, Shabalina SA, Robert C, Fournous G, La Scola B, Raoult D, Koonin EV.** 2011. Viruses with more than 1,000 genes: Mamavirus, a new Acanthamoeba polyphaga mimivirus strain, and reannotation of Mimivirus genes. Genome Biol Evol **3:**737-742.
8. **Yoosuf N, Yutin N, Colson P, Shabalina SA, Pagnier I, Robert C, Azza S, Klose T, Wong J, Rossmann MG, La Scola B, Raoult D, Koonin EV.** 2012. Related giant viruses in distant locations and different habitats: *Acanthamoeba polyphaga* moumouvirus

represents a third lineage of the *Mimiviridae* that is close to the megavirus lineage. Genome Biol Evol **4:**1324-1330.

9. **Van Etten JL, Gurnon JR, Yanai-Balser GM, Dunigan DD, Graves MV.** 2010. Chlorella viruses encode most, if not all, of the machinery to glycosylate their glycoproteins independent of the endoplasmic reticulum and Golgi. Biochim Biophys Acta **1800:**152-159.
10. **Wang IN, Li Y, Que Q, Bhattacharya M, Lane LC, Chaney WG, Van Etten JL.** 1993. Evidence for virus-encoded glycosylation specificity. Proc Natl Acad Sci U S A **90:**3840-3844.
11. **Legendre M, Santini S, Rico A, Abergel C, Claverie JM.** 2011. Breaking the 1000-gene barrier for Mimivirus using ultra-deep genome and transcriptome sequencing. Virol J **8:**99.
12. **La Scola B, Desnues C, Pagnier I, Robert C, Barrassi L, Fournous G, Merchat M, Suzan-Monti M, Forterre P, Koonin E, Raoult D.** 2008. The virophage as a unique parasite of the giant mimivirus. Nature **455:**100-104.
13. **Luther KB, Hulsmeier AJ, Schegg B, Deuber SA, Raoult D, Hennet T.** 2011. Mimivirus collagen is modified by bifunctional lysyl hydroxylase and glycosyltransferase enzyme. J Biol Chem **286:**43701-43709.
14. **Xiao C, Chipman PR, Battisti AJ, Bowman VD, Renesto P, Raoult D, Rossmann MG.** 2005. Cryo-electron microscopy of the giant Mimivirus. J Mol Biol **353:**493-496.
15. **Kuznetsov YG, Xiao C, Sun S, Raoult D, Rossmann M, McPherson A.** 2010. Atomic force microscopy investigation of the giant mimivirus. Virology **404:**127-137.

16. **Boughalmi M, Saadi H, Pagnier I, Colson P, Fournous G, Raoult D, La Scola B.** 2012. High-throughput isolation of giant viruses of the Mimiviridae and Marseilleviridae families in the Tunisian environment. Environ Microbiol**:**DOI: 10.1111/1462-2920.12068.

17. **La Scola B, Campocasso A, N'Dong R, Fournous G, Barrassi L, Flaudrops C, Raoult D.** 2010. Tentative characterization of new environmental giant viruses by MALDI-TOF mass spectrometry. Intervirology **53:**344-353.

18. **Ghigo E, Kartenbeck J, Lien P, Pelkmans L, Capo C, Mege JL, Raoult D.** 2008. Ameobal pathogen mimivirus infects macrophages through phagocytosis. PLoS Pathog **4:**e1000087. doi:1000010.1001371/journal.ppat.1000087.

19. **Shin BK, Wang H, Yim AM, Le Naour F, Brichory F, Jang JH, Zhao R, Puravs E, Tra J, Michael CW, Misek DE, Hanash SM.** 2003. Global profiling of the cell surface proteome of cancer cells uncovers an abundance of proteins with chaperone function. J Biol Chem **278:**7607-7616.

20. **Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M.** 2006. In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat Protoc **1:**2856-2860.

21. **Brand DD, Latham KA, Rosloniec EF.** 2007. Collagen-induced arthritis. Nat Protoc **2:**1269-1275.

22. **Schegg B, Hülsmeier AJ, Rutschmann C, Maag C, Hennet T.** 2009. Core glycosylation of collagen is initiated by two beta(1-O)galactosyltransferases. Mol Cell Biol **29:**943-952.

23. **Snir O, Backlund J, Bostrom J, Andersson I, Kihlberg J, Buckner JH, Klareskog L, Holmdahl R, Malmstrom V.** 2012. Multifunctional T cell reactivity with native and glycosylated type II collagen in rheumatoid arthritis. Arthritis Rheum **64:**2482-2488.

24. **Backlund J, Carlsen S, Hoger T, Holm B, Fugger L, Kihlberg J, Burkhardt H, Holmdahl R.** 2002. Predominant selection of T cells specific for the glycosylated collagen type II epitope (263-270) in humanized transgenic mice and in rheumatoid arthritis. Proc Natl Acad Sci USA **99:**9960-9965.

25. **Becher B, Durell BG, Miga AV, Hickey WF, Noelle RJ.** 2001. The clinical course of experimental autoimmune encephalomyelitis and inflammation is controlled by the expression of CD40 within the central nervous system. J Exp Med **193:**967-974.

26. **La Scola B, Marrie TJ, Auffray JP, Raoult D.** 2005. Mimivirus in pneumonia patients. Emerg Infect Dis **11:**449-452.

27. **Vanspauwen MJ, Franssen FM, Raoult D, Wouters EF, Bruggeman CA, Linssen CF.** 2012. Infections with mimivirus in patients with chronic obstructive pulmonary disease. Respir Med **106:**1690-1694.

28. **Yuki N, Susuki K, Koga M, Nishimoto Y, Odaka M, Hirata K, Taguchi K, Miyatake T, Furukawa K, Kobata T, Yamada M.** 2004. Carbohydrate mimicry between human ganglioside GM1 and Campylobacter jejuni lipooligosaccharide causes Guillain-Barre syndrome. Proc Natl Acad Sci U S A **101:**11404-11409.

29. **Pelletier N, Raoult D, La Scola B.** 2009. Specific recognition of the major capsid protein of Acanthamoeba polyphaga mimivirus by sera of patients infected by Francisella tularensis. FEMS Microbiol Lett **297:**117-123.

30. **Jarvinen P, Aho K.** 1994. Twin studies in rheumatic diseases. Semin Arthritis Rheum **24:**19-28.

31. **Svendsen AJ, Holm NV, Kyvik K, Petersen PH, Junker P.** 2002. Relative importance of genetic effects in rheumatoid arthritis: historical cohort study of Danish nationwide twin population. BMJ **324:**264-266.
32. **Lynn AK, Yannas IV, Bonfield W.** 2004. Antigenicity and immunogenicity of collagen. J Biomed Mater Res B Appl Biomater **71:**343-354.
33. **Stuart JM, Townes AS, Kang AH.** 1982. Nature and specificity of the immune response to collagen in type II collagen-induced arthritis in mice. J Clin Invest **69:**673-683.
34. **Keller A, Nesvizhskii AI, Kolker E, Aebersold R.** 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem **74:**5383-5392.
35. **Nesvizhskii AI, Keller A, Kolker E, Aebersold R.** 2003. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem **75:**4646-4658.

451 **Table 1. Mimivirus surface proteins identified by biotinylation.**

ORF	Protein annotation	% coverage
R459	Uncharacterized protein	92%
L724	Uncharacterized protein	69%
L725	Uncharacterized protein	51%
R489	Uncharacterized protein	49%
R714	Uncharacterized protein	42%
L485	Uncharacterized protein	41%
L330	Uncharacterized protein	36%
R305	Uncharacterized protein	36%
R727	Uncharacterized protein	35%
R345	Uncharacterized protein	35%
L53	Uncharacterized protein	33%
R457	Uncharacterized protein	33%
R362	Thioredoxin domain-containing protein	26%
L488	Uncharacterized protein	24%
R346	Uncharacterized protein	24%
L586	Uncharacterized protein	23%
L829	Uncharacterized protein	23%
R623	Uncharacterized protein	22%
L550	Uncharacterized protein	22%
L425	Capsid protein-1	19%
L647	Uncharacterized protein	19%
R306	Uncharacterized protein	19%
L645	Uncharacterized protein	19%

L591	Uncharacterized protein	19%
L719	Uncharacterized protein	18%
R463	Uncharacterized protein	18%
R705	Uncharacterized protein	17%
L585	Uncharacterized protein	17%
L454	Uncharacterized protein	17%
R653	Uncharacterized protein	16%
R610	Uncharacterized protein	14%
L629	Uncharacterized protein	14%
R710	Uncharacterized protein	13%
L399	Uncharacterized protein	12%
R443	Thioredoxin domain-containing protein	12%
R253	Uncharacterized protein	11%
L324	Uncharacterized protein	11%
R307	PP2C-like domain-containing protein	11%
L492	Uncharacterized protein	11%
R596	Probable FAD-linked sulfhydryl oxidase	10%
R135	Putative GMC-type oxido reductase	9.5%
L442	Uncharacterized protein	9.5%
R347	Uncharacterized protein	9.5%
R622	Putative tyrosine-protein phosphatase	9%
L609	Uncharacterized protein	8.6%
L612	Uncharacterized protein	7.5%
L309	Uncharacterized protein	6.2%
R252	Uncharacterized protein	6.1%
R526	Putative alpha/beta hydrolase	6.1%

R692	Uncharacterized protein	4.8%
L236	Uncharacterized protein	4.7%
L448	Uncharacterized protein	4.5%
L264	Uncharacterized WD repeat-containing protein	4.5%
R588	Uncharacterized protein	3.3%
L605	Structural PPIase-like protein	3%
R553	Uncharacterized protein	2.7%
L71	Collagen-like protein 1	2.5%
L357	Uncharacterized protein	2.4%
R643	Uncharacterized protein	2.2%
L397	Uncharacterized protein	2.1%

Mimivirus surface proteins identified by mass spectrometric peptide sequencing after biotinylation and analyzed by Mascot software. Results were validated by Scaffold (version Scaffold_3.4.9, Proteome Software Inc., Portland, OR) and peptide identifications were accepted if they established >80% probability as specified by the Peptide Prophet algorithm (34). Protein identifications were accepted if they could be established at >99% probability and contained at least 2 identified peptides as specified by Protein Prophet algorithm (35).

458

459 **Table 2. Mimivirus proteins recognized by human serum IgG**

ORF	Protein annotation	Healthy subjects					Rheumatoid Arthritis				
L425	Capsid protein	68	601	463	68	387	983	271	258	70	327
R135	Putative GMC-type oxidoreducatse		370	394		153	774	255	163		161
L410	Core protein		223	284		179	685	123	133		97
R345	Putative regulator of chromosome condensation	52	123	145	28	58	350	185	77		217
R349	Uncharacterized protein	28	28	30	26	26	35	26		28	32

460 Scores for viral proteins recognized by 5 healthy subject sera and 5 rheumatoid arthritis sera
461 are listed in columns for each serum tested. Values indicate Mascot scores representing the
462 probability of positive matches for the recognized proteins. Scores above 25 were significant
463 for $p<0.05$.

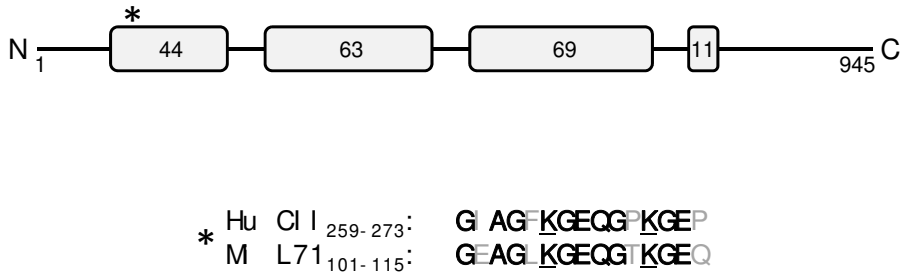

Figure 1. Domain organization of mimivirus L71 protein. The four collagen domains of L71 are shown as grey boxes with the number of G-X-Y repeats given inside. The asterisk shows the position of the sequence motif similar to the epitope human collagen type II recognized as immunodominant in rheumatoid arthritis [24]. The sequence of this human collagen type II (Hu CII) T-cell epitope encompassing amino acids 259-273 is shown aligned with the corresponding sequence of mimivirus L71 (Mi L71) encompassing amino acids 101-115.

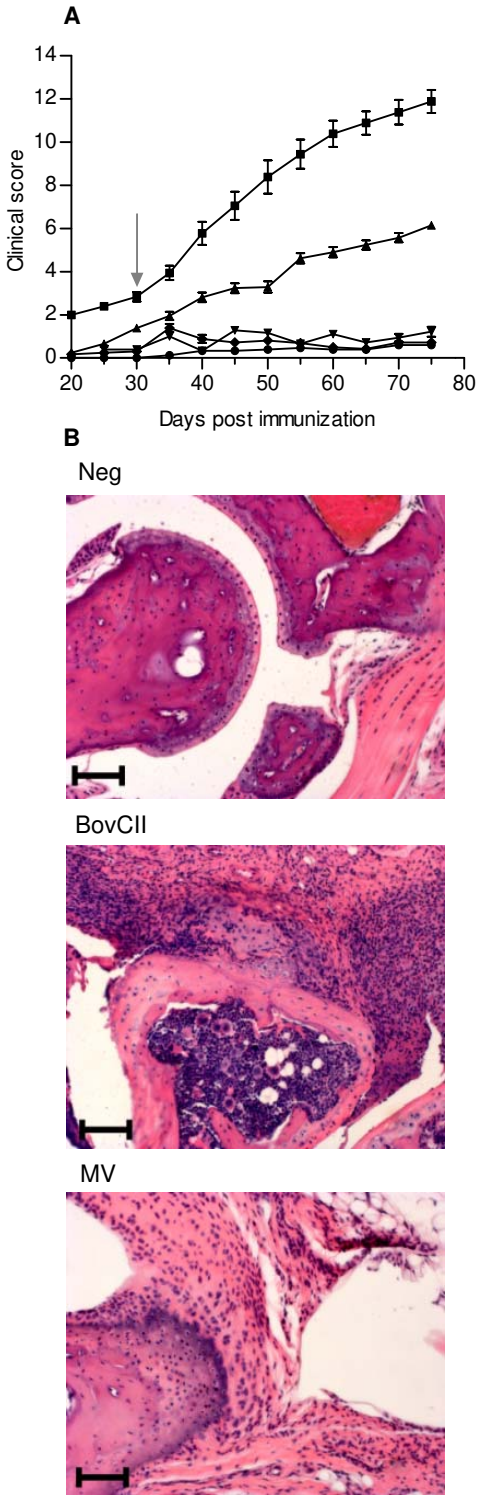

Figure 2. Joint inflammation in DBA/1 mice immunized with mimivirus proteins. (**A**) Clinical severity of arthritic limbs in the groups of PBS (●), bovine collagen type II (■), recombinant L71 protein (▼), marseillevirus proteins (◆) and mimivirus proteins (▲) immunized mice are shown as mean ± SEM. The arrow shows the time point of booster immunization. Data represent three independent experiments including 10-21 mice per group. (**B**) Representative H&E stained sections of hind limbs by day 75 after immunization showing cartilage damage and synovial hyperplasia in bovine collagen type II (BovCII) and mimivirus protein (MV) immunized mice. No sign of pathology were visible in PBS negative control mice (Neg), scale bar 100 µm.

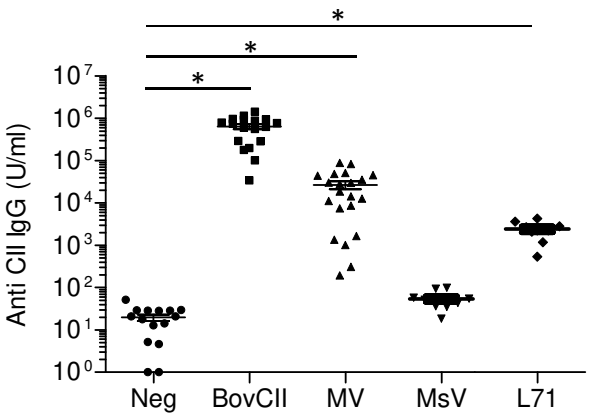

Figure 3. **Anti-collagen type II IgG titers in DBA/1 mice immunized with mimivirus proteins.** Levels of serum IgG measured by ELISA against endogenous mouse collagen type II (CII) in mice immunized with PBS (Neg), bovine collagen type II (BovCII), mimivirus proteins (MV), marseillevirus proteins (MsV), or recombinant L71 protein (L71). Data represent three independent experiments including 10-21 mice per group, horizontal bars show mean ±SEM, *$p<0.01$.

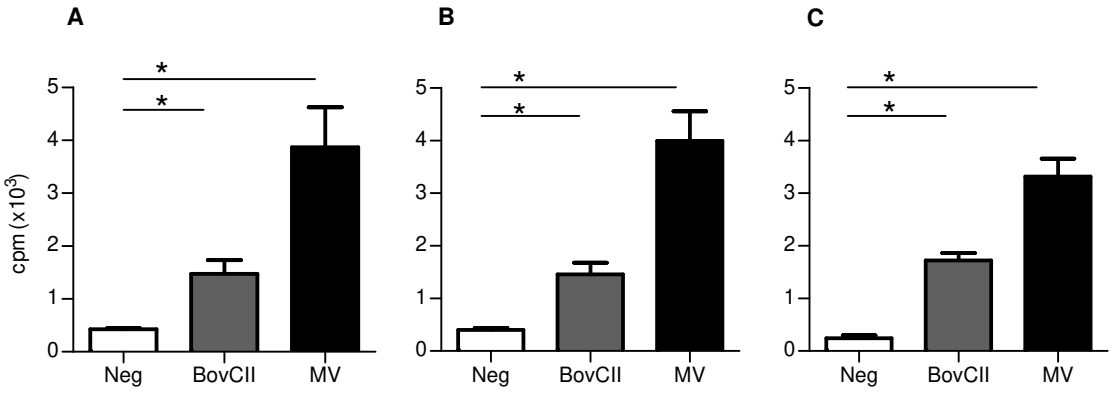

Figure 4. **Auto-reactive T-cell response in DBA/1 mice immunized with mimivirus proteins.** (**A**) Recall responses in cells isolated from draining lymph nodes from mice immunized with PBS (Neg), bovine collagen type II (BovCII), or mimivirus proteins (MV) after stimulation with denatured mouse collagen type II. (**B**) Recall responses after stimulation with denatured bovine collagen type II. (**C**) Recall responses after stimulation with denatured fragmented recombinant mimivirus protein L71. Data represent mean ± SEM from groups of 3 mice, *$p<0.01$.

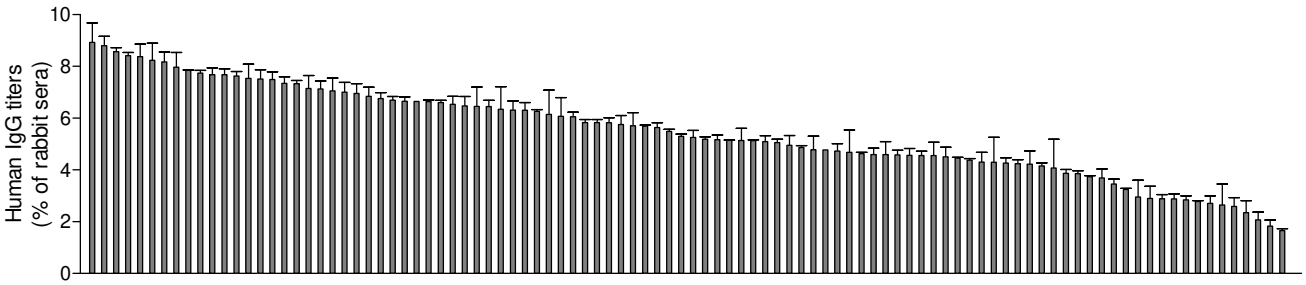

Figure 5. Reactivity of human sera towards mimivirus proteins. Anti-mimivirus IgG titers in sera of 100 healthy subjects were measured by ELISA after dilution 1:100 and expressed as a ratio to IgG titers measured in 1:1000-diluted sera from rabbits previously immunized with mimivirus particles. Data represent mean ± SEM from 4 analyses.

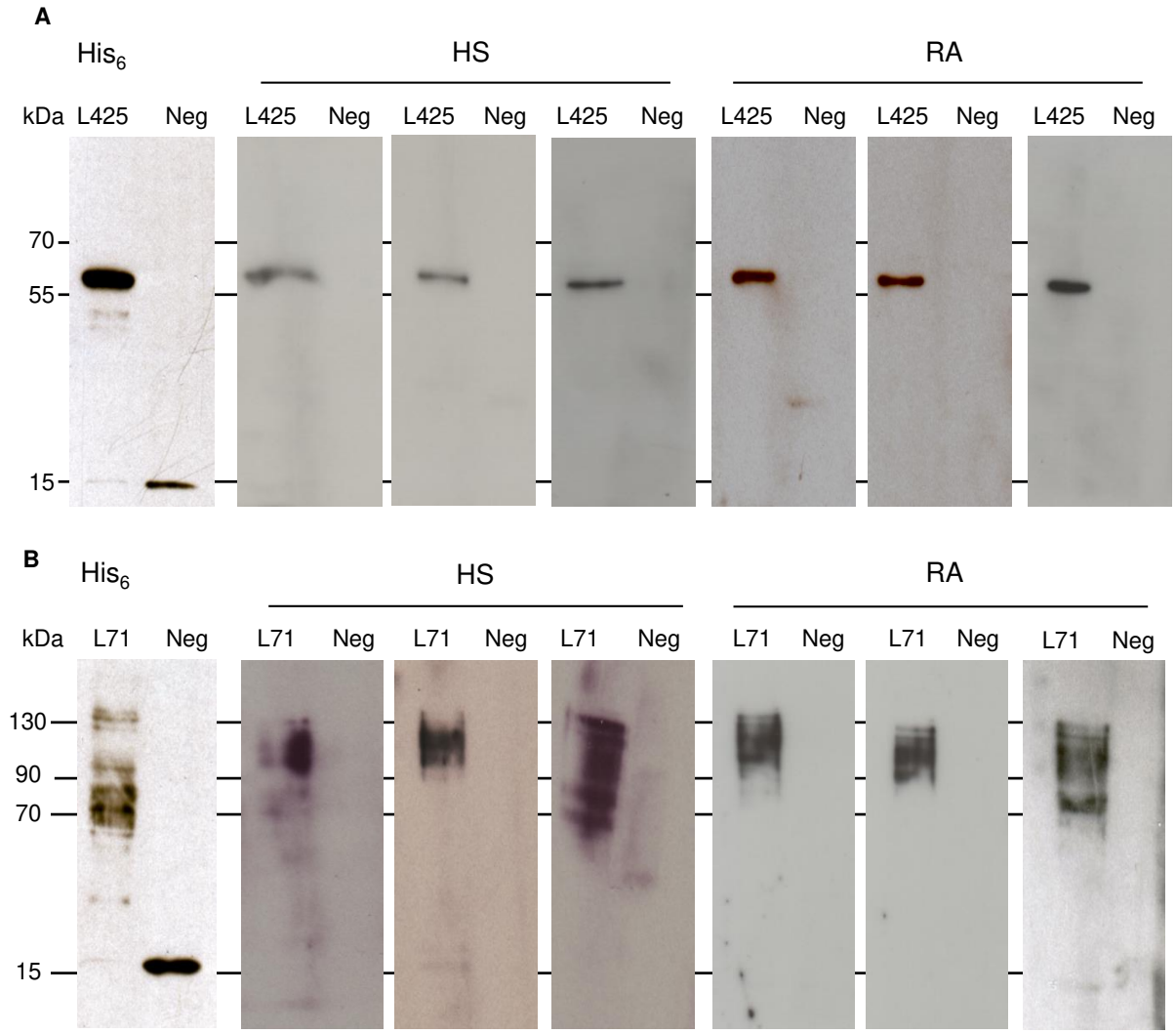

Figure 6. Recognition of mimivirus proteins by human sera. (**A**) Representative Western blots of sera from healthy subjects (HS) and rheumatoid arthritis (RA) patients recognizing mimivirus capsid protein L425. (**B**) Representative Western blots of sera from healthy subjects (HS) and rheumatoid arthritis (RA) patients recognizing mimivirus collagen L71. Sera were diluted 1:4000. Positions of recombinant L425 and L71 proteins in the blots are shown at the left of each panel using an anti-His_6 antibody (His_6). A 15 kDa fragment of His_6-tagged human GLT25D2 protein was used as negative control (Neg).

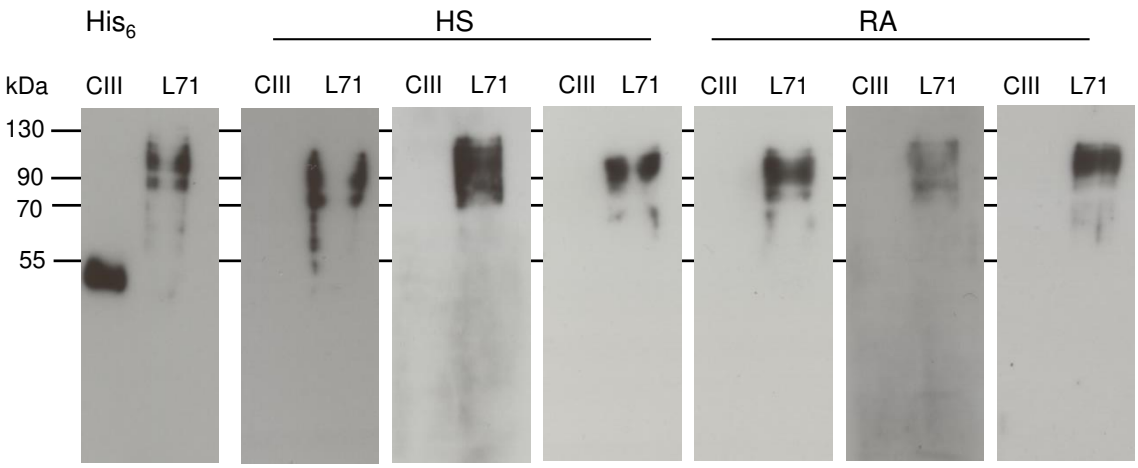

Figure 7. Specific recognition of mimivirus collagen L71 by human sera. Representative Western blots of L71-positive sera from healthy subjects (HS) and rheumatoid arthritis (RA) patients recognizing mimivirus collagen L71 but not a fragment of human collagen type III containing 114 [G-X-Y] repeats (CIII). Sera were diluted 1:4000. Positions of recombinant L71 and CIII proteins in the blots are shown at the left of the panel using an anti-His_6 antibody (His_6).